Discrete Probability

CISC 2210 — CUNY Brooklyn College
Lecture Notes

Discrete Probability: Intro

Probability Theory is the mathematical study of uncertainty and randomness, and it allows us to model and predict the likelihood of various outcomes in uncertain scenarios. **Discrete probability** focuses specifically on systems where the set of possible outcomes is countable or finite, such as rolling dice, flipping coins, or drawing cards. In such discrete contexts, we can list all possible outcomes individually.

This field is especially relevant in computer science, where algorithms often rely on random inputs, simulations, and probabilistic reasoning to function efficiently. Applications of discrete probability in computer science include randomized algorithms, data structures like hash tables, network protocols, artificial intelligence, and machine learning.

Understanding how to compute and manipulate probabilities is crucial for making informed decisions under uncertainty, which is common in automated systems. Moreover, many encryption and cybersecurity techniques rely on probabilistic principles to ensure unpredictability and security of data.

This topic will help us build intuition and formal tools for analyzing randomness in computing contexts.

Visualization of Probability Tools

The study of discrete probability often begins with familiar random devices such as a **six-sided die** and a **two-sided coin**:

- A standard die (plural: dice) is a cube with 6 faces numbered 1 through 6, each assumed equally likely when the die is **fair**, which means we can roll any of these numbers with equal chance.
- A fair coin has two sides — **heads** and **tails** — each occurring with the same chance when the coin is flipped.

These simple tools help us model more complex random systems, including multi-step processes and compound events.

By using these basic tools, we build an intuition for how to define probability models and calculate event probabilities.

In the slides ahead, we'll use dice and coins to illustrate essential concepts like experiments, outcomes, sample spaces, and events.

Visualization of Probability Tools

A 6-sided die. This work is based on "Dice" by Rosa studio licensed under CC-BY-4.0.

A coin (dime). This work is based on "Mercyry Dime" by Eugene Korolev licensed under CC-BY-4.0.

Basic Concepts

An **experiment** in probability is a process or action with uncertain results, whose outcomes can be observed or measured.

An **outcome** is a single possible result of an experiment, such as rolling a 4 on a die or getting heads on a coin flip.

The **sample space**, denoted Ω (`$\Omega$`: the Greek capital letter "Omega"), is the set of all possible outcomes of an experiment; for example, for a die roll, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

For a coin flip, $\Omega = \{\mathrm{H}, \mathrm{T}\}$, and for flipping two coins, $\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$.

Experiments must be clearly defined to properly construct sample spaces and ensure accurate probability modeling.

Sample spaces may be finite, countably infinite (similar to the infinity of $\mathbb{N}$: the set of integers,) or uncountable (similar to the vast infinity of $\mathbb{R}$: the set of real numbers,) but in discrete probability, we focus on the first two types.

Basic Concepts

An **event** is any subset of the sample space and represents a set of outcomes that share some common feature. Events can be simple (containing one outcome) or compound (containing multiple outcomes).

For example, in a single die roll, the event “roll an even number” is $E = \{2, 4, 6\}$.

The **probability** of an event A, written $P(A)$ (`$P(A)$`) or $\text{Pr}[A]$ (`$\text{Pr}[A]$`), is the sum of the probabilities of the outcomes in A.

If all outcomes in the sample space are equally likely, then $P(A) = \frac{\text{number of outcomes in } A}{\text{total number of outcomes in } \Omega}$.

Example: You roll a fair die and want to find the probability of rolling a number greater than 4. The event is $\{5, 6\}$, so we compute that $P(\{5, 6\}) = \frac{2}{6} = \frac{1}{3}$. In other words, the chance that you roll a number greater than 4 is $33.\overline{33}\%$.

Basic Concepts

Any questions so far?

Basic Concepts

A **probability model** consists of:

1. A sample space Ω,
2. A collection of events (subsets of Ω), and
3. A function P, called the **probability measure**, that assigning probabilities to events.

The function P maps each event to a number in the interval $[0, 1]$ (real numbers from 0 to 1, inclusively), where $P(\Omega) = 1$.

In the uniform model, all outcomes are equally likely, and the probability of any event is computed by counting favorable and total outcomes. In more general, real-life models, however, probabilities can differ from outcome to outcome; for example, in a biased die, some faces might occur more often than others.

We use these models to predict the likelihood of events and simulate randomness in a mathematically rigorous way.

Basic Concepts

The foundations of probability theory rest on three basic rules known as **Kolmogorov's axioms**:

1. For any event A, the probability satisfies $0 \leq P(A) \leq 1$, meaning it is a real number between 0 and 1.
2. We previously stated that: $P(\Omega) = 1$, which reflects that the total probability over the entire sample space must equal 1.
3. If two events A and B are mutually exclusive (i.e., $A \cap B = \emptyset$), then $P(A \cup B) = P(A) + P(B)$. Otherwise, if they aren't mutually exclusive, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

These axioms allow us to use familiar set operations such as **unions**, **intersections**, and **complements** to reason about events.

For example, the event "not rolling a 6" ($E_{\text{Not } 6} = \{1, 2, 3, 4, 5\}$) is the complement of rolling a 6 ($E_6 = \{6\}$). According to axioms 2 and 3 above, this means that $P(E_6) + P(E_{\text{Not } 6}) = 1$, so $P(E_{\text{Not } 6}) = 1 - P(E_6) = 1 - \frac{1}{6} = \frac{5}{6}$.

Exercises

Fun Class Activities:

1. If you flip a fair coin and roll a die, how many total outcome pairs are there? An example of an outcome pair is $(2, \mathrm{H})$.
2. List all outcomes where the coin shows heads and the die shows an even number.
3. What is the probability of getting a "3" on a fair die and "Tails" on a fair coin in a single trial?
4. You roll 2 fair 6-sided dice at the same time. What is the probability to get a double (= the same number on both dice)? Examples: $(1, 1)$, $(5, 5)$, etc.
5. A coin has $P(\mathrm{H}) = 0.6$ and $P(\mathrm{T}) = 0.4$. What is the probability of getting 2 heads in 2 independent flips? ***Hint***: Use the multiplication rule.
6. Let $\Omega = a, b, c$ for some experiment with $P(a) = 0.3$ and $P(b) = 0.5$. What is $P(c)$? ***Hint***: Refer to the axioms.
7. If you roll a fair die, what is the probability that you roll a number less than 5?
8. If $P(A) = 0.7$, what is $P(A')$, the probability of the complement of A?

Exercises

Any questions so far?

Basic Concepts

A few more rules can be inferred from the three axioms:

4. $P(\emptyset) = 0$
5. $P(A^c) = 1 - P(A)$ [A^c is the complement of A.]
6. If $A \subseteq B$, then $P(A) \leq P(B)$

Examples:

- The probability of rolling a 7 on a die is 0 since rolling a 7 is a non-existent (impossible) outcome.
- The probability of rolling anything but 4 on a die is $\frac{5}{6}$ since rolling a 4 has $P[\text{rolling a } 4] = \frac{1}{6}$, so $P[\text{not rolling a } 4] = 1 - P[\text{rolling a } 4] = 1 - \frac{1}{6} = \frac{5}{6}$.
- The event of "rolling an even integer" is a subset of the event "rolling a number bigger than 1" (**why so?**). Indeed, we have $P[\text{rolling an even integer}] = \frac{3}{6} = \frac{1}{2}$ and $P[\text{rolling a number bigger than } 1] = \frac{5}{6}$, and that $\frac{3}{6} < \frac{5}{6}$.

Conditional Probability

Conditional probability is the probability that an event A occurs given that another event B has already occurred.

It is defined as $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ (`$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$`), provided that $P(B) > 0$.

This formula reflects how new information about one event can affect our belief in another event's occurrence.

For example, suppose we draw a card from a deck and we know it is a red card; because 26 of the cards in the deck are red (the rest are black,) the probability that the drawn card is a heart is $P(\text{heart} \mid \text{The card is red}) = \frac{P(\text{heart} \cap \text{The card is red})}{P(\text{The card is red})}$ $= \frac{P(\text{heart})}{P(\text{The card is red})} = \frac{13/52}{26/52} = \frac{13}{26} = \frac{1}{2}$.

Conditional probability is useful in sequential experiments, decision-making, and updating probabilities based on evidence. It will also help us understanding independence and Bayes' theorem, both of which are crucial for probabilistic reasoning.

Conditional Probability

More examples:

- A card is drawn at random from a standard 52-card deck. Let event A be "the card is a queen" and the event B be "the card is a face card" (there are 12 face cards: King, Queen, and Jack $\times 4$, and 4 queens.) Find $P(A \mid B)$.

  Solution: $P(A \mid B) = P(\text{queen} \mid \text{face}) = \frac{P(\text{queen} \cap \text{face})}{P(\text{face})} = \frac{4/52}{12/52} = \frac{4}{12} = \frac{1}{3}$.

- In a certain region, 10% of users run outdated software. Among them, 40% experience a crash. Find the probability a user experiences a crash given they run outdated software.

  Solution: $P(\text{crash} \mid \text{outdated}) = \frac{P(\text{crash} \cap \text{outdated})}{P(\text{outdated})} = \frac{0.4 \times 0.1}{0.1} = \frac{4/100}{1/10} = \frac{4}{10}$.

- In a dataset, 30% of emails are spam. Among spam emails, 80% contain the word "free". Find the probability that an email contains "free" given that it is spam.

  Solution: $P(\text{contains "free"} \mid \text{spam}) = \frac{P(\text{contains "free"} \cap \text{spam})}{P(\text{spam})} = \frac{0.8 \times 0.3}{0.3} = \frac{0.24}{0.3} = 0.8$.

Independence and Mutual Exclusivity

Two events A and B are called **independent** if the occurrence of one does not affect the probability of the other, which means $P(A \mid B) = P(A)$.

This is equivalent to saying $P(A \cap B) = P(A) \cdot P(B)$, a key identity used frequently in both theory and application.

For example, flipping a coin and rolling a die are independent experiments because they don't influence one another's outcome.

In contrast, two events are **mutually exclusive** if they cannot occur at the same time, meaning $A \cap B = \emptyset$, and thus $P(A \cap B) = 0$.

A roll of 2 and a roll of 3 on a single die are mutually exclusive since both outcomes cannot occur on a single trial.

Importantly, independence and mutual exclusivity are not the same; mutually exclusive events are always dependent because knowing one occurred changes the probability of the other to zero.

Independence and Mutual Exclusivity

More examples:

- Two dice are rolled. Let event A be "first die shows even," and B be "second die shows a 5." Are these events independent?
  Solution: Yes, because $P(\text{'first: even'} \cap \text{'second: 5'}) = P(\text{'first: even'}) \cdot P(\text{'second: 5'}) \ = \frac{3}{6} \cdot \frac{1}{6} \ = \frac{1}{12}$.
- A single die is rolled once. For that single rolling, let A be "roll a 2," and B be "roll a 5." Are these events independent?
  Solution: No, because $P(A \cap B) = 0$, but $P(A) \cdot P(B) \ = \frac{1}{6} \cdot \frac{1}{6} \ = \frac{1}{36}$, so $P(A \cap B) \ \neq P(A) \cdot P(B)$. In fact, these events are mutually exclusive.

Bottom line: Two independent events are never mutual exclusive, and two mutual exclusive events are never independent.

Bayes' Formula

Bayes' formula, named after the 18th century's English statistician Thomas Bayes, is a powerful tool that allows us to reverse conditional probabilities when direct computation is difficult or impossible. It plays a central role in statistical inference, allowing us to update our beliefs about the likelihood of a cause given some observed effect.

The formula states that $P(A \mid B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)}$, where A is a hypothesis and B is observed evidence.

For example, in email filtering, if A is "message is spam" and B is "message contains the word 'damn'," Bayes' rule helps estimate the probability that the message is spam given that the word 'damn' is in the message.

Another application is in medical testing, where the base rate of a disease, false positive rate, and true positive rate can be used to compute the actual chance of illness after a positive test. Moreover, Bayes' theorem bridges observed data and theoretical probabilities, making it foundational to machine learning, decision theory, and artificial intelligence.

Bayes' Formula

<u>More examples</u>:

- In a certain country, 1% of the population has a disease. The test has 99% sensitivity and 95% specificity. Let D: "person has disease" and T: "test is positive". What percentage of people has the disease given that the test turned out positive?

  Solution: $P(D \mid T) = \frac{P(T|D)P(D)}{P(T|D)P(D)+P(T|D')P(D')} = \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01+0.05 \cdot 0.99} \approx 0.166.$
- In an email server, 40% of emails are spam. Among spam, 30% include "win". Among non-spam, $5\%" include "win". Given "win" is in the email, what's the probability that it is spam?

  Solution: $P(\text{spam} \mid \text{contains "win"}) = \frac{0.30 \cdot 0.40}{0.30 \cdot 0.40+0.05 \cdot 0.60} \approx 0.8.$
- Bag A has 2 red and 3 blue marbles. Bag B has 4 red and 1 blue marbles. One bag is chosen at random, and a red marble is drawn. What is the probability that it came from bag A?

  Solution:

  $$P(\text{bag } A \mid \text{red}) = \frac{P(\text{red}|\text{bag } A)P(\text{bag } A)}{P(\text{red}|\text{bag } A)P(\text{bag } A)+P(\text{red}|\text{not bag } A)P(\text{not bag } A)} = \frac{(2/5)(1/2)}{(2/5)(1/2)+(4/5)(1/2)} = \frac{2}{6} = \frac{1}{3}$$

  .

Exercises

Fun Class Activities:

1. A card is drawn from a standard deck. What is the probability that it's a 2, given that it's black?
2. If $P(A \cap B) = 0.15$, and $P(B) = 0.3$, compute $P(A|B)$.
3. If $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$, are A and B independent?
4. If two events are mutually exclusive, what is $P(A \cap B)$? Why?
5. If $P(A) = 0.2$, $P(B|A) = 0.6$, $P(B|A^c) = 0.3$, and $P(A^c) = 0.8$, find $P(A|B)$.

Exercises

Any questions?

Discrete Probability: Sources

Aspnes, James. (2022). Chapter 12. *Notes on Discrete Mathematics*. License: CC BY-SA: Attribution-ShareAlike.

"CS202: Discrete Structures | Unit 5." *Saylor Academy*. License: CC BY: Attribution.

Davis, Stephen. (2019). Chapter 4. *A Cool Brisk Walk Through Discrete Mathematics*, version 2.2. License: CC BY-SA: Attribution-ShareAlike.

Keller, Mitchel T and Trotter, William T. Chapter 10. *Applied Combinatorics*. Open Textbook Library, 2017. License: CC BY-SA: Attribution-ShareAlike.

Lumen Learning. College Algebra | Probability. License: CC BY: Attribution.